10.1098/rsta.2003.1187

THE ROYAL
SOCIETY

Self-organization in evolution: a mathematical perspective

By Ian Stewart

Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK

The neo-Darwinian view of evolution centres upon the role of the gene. Here there seems to be little scope for self-organization. This conclusion is reinforced by traditional models of polymorphism in terms of allele frequencies in a mean-field gene-pool. However, models based on phenotypes, and including nonlinear and collective effects, suggest that evolution can indeed be viewed as a process whereby the ecosystem self-organizes. Here we focus on the phenomenon of speciation, and discuss a series of phenotypic models which together illuminate some of the issues surrounding the role of self-organization, including new approaches to fitness landscapes and species selection. All of these models represent speciation as a symmetry-breaking bifurcation, but in different mathematical contexts including deterministic dynamical systems, stochastic dynamical systems, and iterated function schemes. The main conclusions are surprisingly robust, despite the diversity of the models.

Keywords: evolution; speciation; network; symmetry-breaking

1. Introduction

To what extent can evolution be considered as a self-organizing process? If it can, what is the ‘self’ that does the organizing? And how does the organization process work?

In 1859, Charles Darwin published his *Origin of species* (Darwin 1859), possibly the most influential scientific work of all time. Its central thesis can be restated in today’s jargon: just as human beings can organize systematic changes in living creatures by ‘unnatural selection’, so the creatures themselves can organize such changes by natural selection. As is well known, Darwin did not use the term ‘evolution’, although he did use ‘evolved’ in his closing sentence. But the concept of natural selection was the key step to evolutionary biology.

It is arguable that all allegedly self-organizing systems are organized not just by themselves, but by their contexts. Even the Universe as a whole is organized in the context of the laws of physics. The phrase ‘self-organizing’ should not be taken too literally. For example, in this Darwinian vision, the organism does not organize itself. The process of natural selection requires the active intervention of other organisms, such as predators or parasites, and the environment also plays a significant role. In contrast, it is more reasonable to see the ecosystem as self-organizing—again,

One contribution of 18 to a Theme ‘Self-organization: the quest for the origin and evolution of structure’.

remembering the constraints imposed by context. So by 'self-organizing' we mean that, in response to the appropriate context, the system becomes organized as a consequence of its own dynamic, rather than organization being imposed directly by some external influence.

Today, Darwin's point of view has been all but replaced by a more specific, reductionist stance: the neo-Darwinian contention that the 'unit of selection' is the gene. Since genetic mutations and recombinations are random, this stance seems to leave little room for self-organization. We shall argue that this appearance is deceptive, and that the neo-Darwinian position is by no means inconsistent with a kind of self-organization: not of individual organisms, but on a collective level.

Our arguments rest on simple mathematical models of evolutionary processes. The purpose of these models is not to represent biological evolution in a realistic way—the process is far too complex for that. It is to consider certain features that the evolutionary process may or may not possess, to clarify the relationships between those features, and to enquire whether any of them correspond to features observed in the real world. This is a reasonable role for mathematics: not to assert what *is*, but to consider what *might be*. By eliminating what might be, we can hope to emulate Sherlock Holmes and discover what *is*.

Darwin's description of the evolutionary process is purely verbal. The best-known mathematical formalism, classical population genetics, goes back to Fisher (1930). In his approach, there is no mathematical correlate of an organism as such; instead, the model operates in terms of allele frequencies in a mean-field gene-pool. In keeping with the mathematical techniques available in Fisher's day, the theory is essentially a linear one (effects are proportional to causes). Such an approach can be successful (and was) in well-defined and localized systems, such as plant-breeding; however, these successes do not necessarily translate into the complexities of naturally occurring ecosystems.

Because it is mean field and linear, classical population genetics is severely limited in the range of phenomena that it can address. In particular speciation (the formation of new species) is somewhat alien to Fisher's viewpoint. The closest analogue, the formation of 'stable polymorphisms' (Maynard Smith 1982), is viewed as a somewhat rare circumstance. However, it is relatively common in the real world.

Biologically, we cannot observe evolution in any detail on megayear time-scales, and there is only a small number of field studies of the effects of selection on time-scales of decades, notably the work of Grant and co-workers (Grant *et al.* 1976; Grant 1991; Grant & Grant 2002). But from a mathematical perspective, some conclusions of Fisher's mathematical models seem to reflect built-in limitations of classical population genetics, rather than innate features of the real world. Fisher's linear stance renders his models insensitive to nonlinear effects, and his reliance on mean-field allele proportions renders them insensitive to collective effects in groups of organisms.

We shall describe several models that attempt to address these issues of nonlinearity and collective behaviour. Technically, the models are dynamical systems—collections of ordinary differential equations (ODEs). However, the equations have a special feature, namely, a network structure. They are examples of what might be called *network dynamics*, and the network topology is reflected in the form of the equations. In spirit, therefore, they fall into the realm of complex adaptive systems (Auyang 1998; Kauffman 1993, 1995; Mainzer 1994).

Unlike most current evolutionary models, the ones considered here focus on phenotype, not genotype. It would not be difficult, in principle, to include genetic effects in the models, and the main conclusions would probably survive this modification, but we will not pursue that path here. The BirdSym simulation of Elmhirst (2001) is an example of what can be done. Our stance here will be that the dynamic of evolution is principally driven by selection, and that this operates on individual organisms—in the context of other organisms and the overall environment. The role of the gene is to render the potential phenotype 'fluid', to provide a range of phenotypes from which natural selection can *select*. What we have in mind here is not mutation, the effects of which are usually cryptic and operate over very long time-scales, but recombination, which produces substantial genetic diversity in each new generation. Thus we are thinking of sexual species, even though the model does not make any explicit distinction between the sexes. The reasons for focusing on recombination, and the probable genetic context in which our phenotypic models operate, are discussed in Cohen & Stewart (2000) and Stewart *et al.* (2003).

There is a close relationship between our approach and that of 'adaptive dynamics' (Kisdi & Geritz 1999). The main differences are that we emphasize phenotype, employ model ODEs that are structured as complex systems, and focus on speciation in homogeneous environments, whereas adaptive dynamics centres on genotype, uses ODEs based on allele frequencies and considers heterogeneous environments. The underlying philosophies, especially the central role afforded to selection, are very similar, and the two approaches complement each other. With further work it should be possible to combine them.

There are many theories of speciation, and many scales on which speciation-like phenomena can occur (see Rice & Hostert 1993). For convenience we will use the term 'speciation' to denote any of these; in particular we do not distinguish between large-scale speciation and smaller-scale polymorphism. Our models probably correspond more closely to polymorphism, but a lot depends on the scales of population and time appropriate to the type of speciation event, and the amount of fine detail ignored in the model. In other words, the model should be considered as a broad guide, and should not be taken too literally.

Biological theories of speciation fall into two main classes: allopatric and sympatric. In allopatric speciation (Mayr 1963, 1970), a (small) founder population becomes geographically separated from the main group, evolves independently in isolation and changes so much that, when it finally rejoins the main group, its members cannot (or, perhaps more importantly, *do* not) interbreed with that group. This is one way, the most obvious way, to ensure that the stabilizing (that is, phenotypically unifying) effects of 'gene flow' among interbreeding organisms is nullified.

In sympatric speciation, there is no *obvious* trigger for the separation of the group into two nascent species. The environment is assumed to be homogeneous (though changing over time), and interbreeding is panmictic (all organisms can potentially interbreed). Nevertheless, there are more subtle triggers, and the population can still split into two phenotypically distinct parts. Until recently, sympatric speciation has been thought to be rare, and allopatric speciation commonplace. The recent literature, though, focuses increasingly on sympatric mechanisms, and these now seem to be both natural and widespread (see Kawecki 1997; Dieckmann & Doebeli 1999; Higashi *et al.* 1999; Kondrashov & Kondrashov 1999; Tregenza & Butlin 1999; Wolffe & Matzke 1999; Huey *et al.* 2000; Rundle *et al.* 2000).

The main psychological obstacle to accepting sympatric speciation as a possible mechanism is the question: what, other than geographical isolation, can disrupt gene flow between the nascent species? Why does panmictic interbreeding not reconstitute the original single group? Since the organisms are all evolving towards some nearby peak of the 'fitness landscape', and they are all starting in the same place, why do they not all evolve to the *same* peak?

We will argue that nonlinear and collective effects answer these objections, and therefore should not be ignored in models. Even if the specific models proposed here are wrong (and without further refinement this is likely, since they are advanced in order to analyse the mathematical issues in a simple setting), they reveal how different the picture of speciation becomes as soon as we remove the classical blinkers and model evolution with the tools of the 21st century. And they cast new light on some long-running controversies, such as the status of 'species selection' and the concept of a 'fitness landscape'.

2. Symmetry-breaking bifurcations

A ubiquitous phenomenon in nonlinear dynamics is the occurrence of *bifurcations*: major changes in system behaviour that result from gradual changes in parameters (see Hale & Koçak 1991; Seydel 1994). For example, in steady-state bifurcation, the number of steady states may change, implying that a new steady state has appeared in the system or an old one has disappeared. Such a change is generally associated with a stable steady state becoming unstable. Similarly, in Hopf bifurcation, a stable steady state loses stability (in a different way) and a limit cycle appears, corresponding to time-periodic oscillations.

Speciation is clearly a major, qualitative change in behaviour. If we model evolution using a nonlinear dynamical system, then speciation is likely to correspond to some kind of bifurcation. From a nonlinear perspective, speciation is much less surprising than its complete absence would be. Although this does not *explain* speciation, it warns us to beware of models or theories in which speciation is viewed as an exceptional event happening only in extremely unusual circumstances.

The models that we discuss here interpret speciation as a symmetry-breaking steady-state bifurcation (Golubitsky *et al.* 1988; Golubitsky & Stewart 2002*a*). Here not only the number of steady states, but also their symmetry properties, change. The underlying intuition is that an evolutionary state corresponding to a single species is (at least approximately) symmetric under all permutations (or relabellings) of the organisms concerned, whereas a state corresponding to several species has less symmetry because permutations that mix up distinct species do not preserve the dynamic. The original motivation for the models was to formalize this idea.

In this section we introduce the simplest type of dynamical model that represents speciation as a symmetry-breaking bifurcation, describe its implications and compare these with some limited observational information. Later sections take the analysis further and investigate how the model can be modified to make it more realistic.

(*a*) *Idealized model*

We begin with an idealized model, which illustrates some key ideas but is simple enough to be analysed using standard methods of nonlinear dynamics. It was introduced in Cohen & Stewart (2000) and studied in Stewart *et al.* (2003), Elmhirst

(1998, 2001) and Dias & Stewart (2003). In later sections we discuss modifications that render the model more 'realistic', and show that the main conclusions remain valid. All this suggests that speciation phenomena are fairly robust, even in the sympatric case, with many different models leading to broadly similar conclusions. (In the allopatric case, the robustness of speciation is far more intuitive. Allopatric versions of the model can easily be constructed, and the resulting behaviour seems to be broadly consistent with this belief.) From this perspective, the classical models appear to form an unrepresentative and limited class.

There are several ways to motivate this model. The approach that we now explain differs slightly from earlier work, but has some technical advantages.

We represent the phenotype of an organism by a vector $x = (x_1, \dots, x_k)$ in some *phenotypic space* $X = \mathbb{R}^k$. For example, the organisms might be birds, with x_1 the beak length; x_2, weight; x_3, wingspan; x_4, ability to digest cactus; and so on. Individual i has phenotype

$$x^i = (x^i_1, \dots, x^i_k) \in \mathbb{R}^k.$$

In order for this representation to be associated with a species, we assume that the variables are observed only for some fixed age class (and sex) in the overall population.

A population of organisms defines a cloud of points in phenotypic space, with one point per organism. From the viewpoint of complex systems it makes sense to model such a cloud as an explicit collection of individual entities. The entities might represent organisms, or typical samples of organisms, or coarse-grained 'tokens' for organisms. In such models, we can hope to observe collective effects that are invisible to a mean-field approach.

Speciation is a feature of the geometry of clouds in phenotypic space: a kind of phenotypic clumping. If the organisms belong to a single species, then they will be phenotypically similar, and the cloud will form a relatively tight clump. In the idealized limit of identical organisms, the clump will reduce to a single point. If there are several species, the cloud will divide into several clumps (or several points in the idealized limit). From this point of view, speciation is a change from a single clump to several clumps.

Figure 1 illustrates incipient splitting into two clumps in the horned beetle *Podischnus agenor*. Here the chosen phenotypic space has two dimensions: length of head and width of wing cases. The points are unevenly clumped, and seem close to dividing into two distinct clumps. The corresponding beetles have different behaviours: the large 'majors' employ their horns to fight for mating sites, the smaller 'minors' do not. This is not (yet) speciation, but it could be a possible precursor.

The above representation is cumbersome from a modelling point of view: points disappear when the corresponding organism dies or moves out of the chosen age class, and new points appear when new organisms are born and attain the required age. These effects can be modelled using a stochastic process; alternatively, they can be described by a continuum approximation (valid for large enough populations), in which the cloud is represented by a probability density function (PDF) $p = p(x, t)$ on X. We follow the second approach here. The probability that at time t a randomly selected organism (of the right age and sex) has phenotype in some subset $Y \subseteq X$ is

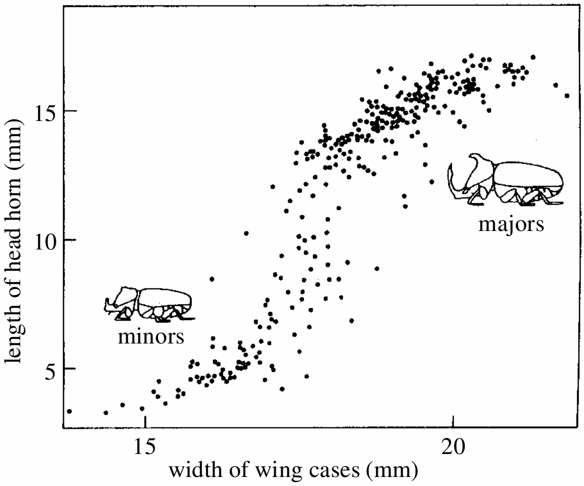

Figure 1. Phenotypic-space representation of a population of horned beetles. Reproduced with permission from Maynard Smith (1982).

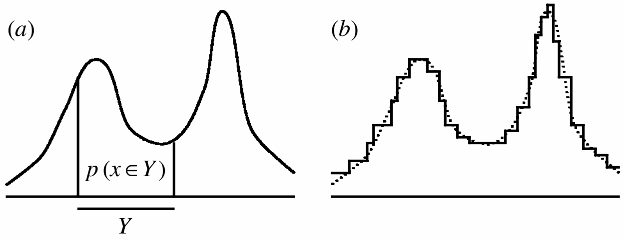

Figure 2. (*a*) Probability density function; (*b*) approximation by a sum of step-functions.

given by the integral

$$p(x \in Y \text{ at time } t) = \int_Y p(x,t)\,\mathrm{d}x.$$

Figure 2*a* illustrates this in the case when there is a single phenotypic variable $(k = 1)$.

We obtain our model by discretizing p, that is, by coarse-graining the population. The coarse-grained distribution of phenotypes can be defined in terms of N 'tokens', which represent a typical sample of organisms from the population. We therefore choose N organisms at random, observe their phenotypes x^i $(1 \leqslant i \leqslant N)$, define a step-function which takes value $1/Nw$ on an interval of length w centred on x^i, and sum all of these step functions as i runs from 1 to N (see figure 2*b*). We use step-functions for programming convenience; Elmhirst (2001) used Gaussians.

It may seem perverse to begin with a discrete description, replace it by a smoothed continuum model and then discretize the result. However, such procedures are commonplace. For example, in elasticity theory, the atomic structure of a metal rod is replaced by a smooth continuum, and the resulting equations are solved by finite-element methods. Our approach here is analogous, and is used for similar reasons.

We assume that these movements can be represented by an ODE. Biologically, this assumption requires the tokens to be chosen in such a way that the labels i

are assigned consistently as time passes, for example, by always choosing the ith token to represent the same lineage. It also assumes a degree of smoothness in the evolutionary process, on the appropriate scales of population size and time.

If this assumption is justified, then the changes to the token phenotypes can be represented by equations of the form

$$\frac{\mathrm{d}x^i}{\mathrm{d}t} = f_i(x^1, \dots, x^N, a), \tag{2.1}$$

where a is a vector of environmental parameters (such as climate, levels of various resources, and so on). The functions f_i will be left unspecified for the moment, but they would represent the 'direction' of selection pressure on a given phenotype, in the context of the current distribution of phenotypes and for environment given by the current value of a. We will see that for conceptual purposes, it is not the specific form of the f_i, but their general structure, that is crucial.

In Cohen & Stewart (2000) the model (2.1) is motivated rather differently: each x^i represents a placeholder for organism dynamics (POD), which is a coarse-grained lineage of representatives of the population. We suspect that the approach through a PDF is technically superior, because it explains the role of N and the x^i in a mathematical rather than a biological manner. Indeed, we conjecture that there is some integro-partial differential equation for p, of which (2.1) is a discretization, and that in a sense this is the real model that should be used, and the ODE structure will be a simple consequence. However, space forbids further discussion of this idea here.

(*b*) *Symmetry of the model*

Another advantage of the approach via a PDF is that it makes plain a natural symmetry constraint on (2.1), which is central to our viewpoint. This constraint is that the system (2.1) should not depend on the choice of labellings of the tokens x^j. The sum of the corresponding step functions is the same if the labels are permuted. This 'shuffling' symmetry implies that the functions f_i should be *equivariant* under the action of the symmetric group $\boldsymbol{S}_N$ of all permutations of N. More precisely, if $\sigma \in \boldsymbol{S}_N$, then

$$f_{\sigma(i)}(x^1, \dots, x^N) = f_i(x^{\sigma^{-1}(1)}, \dots, x^{\sigma^{-1}(N)}) \tag{2.2}$$

for all $\sigma \in \boldsymbol{S}_N$, $x \in X$.

Condition (2.2) distinguishes our approach from most other models in the literature, a notable exception being Vincent & Vincent (2000). Symmetry is a natural concept in connection with speciation—it is intimately associated with the most obvious property of a species, that phenotypes are very similar. The mathematical effects of symmetry (even if only approximate) on dynamics are substantial (Golubitsky *et al.* 1988; Golubitsky & Stewart 2002*a*), and any model that ignores symmetry is likely to miss important collective phenomena. Symmetry is also natural in complex systems, because of the use of identical rules for entities; it plays a central role in pattern formation in complex systems, although that role is seldom made explicit. We will shortly see that generic bifurcation in $\boldsymbol{S}_N$-symmetric systems differs in important ways from generic bifurcation in asymmetric systems (such as classical niche models of character displacement (Salthe 1972) or Fisher-style mean-field genetics).

Condition (2.2) also has a reasonable biological interpretation, which is how it was derived in Cohen & Stewart (2000). Namely, it expresses the fact that initially we have a single species of nominally identical PODs. We then have the following arguments.

(i) POD number 1, say, can distinguish itself from the others, but all the others appear effectively identical to it. Its interactions, in competition for food or mates, will be the same if the other PODs are relabelled. However, the interactions within a POD will normally differ from those between distinct PODs. (If in fact the interactions are the same in both cases, this merely leads to a special case of the same model, and does not alter any significant conclusions.)

(ii) Every POD is in the same position as POD 1, as regards its interactions with itself and with other PODs.

These conditions do *not* imply that all PODs behave identically. In fact, the first condition implies that f_i should be invariant under all permutations of labels $1, 2, \dots, i-1, i+1, \dots, N$, that is, of all labels except label i. The second implies that f_j should be obtained from f_i by interchanging labels i and j. It turns out that these two statements are mathematically equivalent to condition (2.2). Moreover, *without* being more specific about f, we can already deduce some interesting consequences from these natural symmetry constraints. The most fundamental of these is symmetry breaking.

To make these deductions, we consider the equilibria of (2.1), given by

$$f_i(x) = 0, \quad 1 \leqslant i \leqslant N,$$

and interpret these equilibria as distributions of phenotypes. If the functions f_i are chosen to be linear, then there is only one equilibrium, namely, the origin. This equilibrium may be stable or unstable; if it is unstable, then the phenotypes diverge exponentially to infinity and the model is useless; if it is stable, then speciation is impossible. However, these conclusions are (typical) artefacts of the choice of a linear model. When the f_i are nonlinear, far more interesting phenomena occur, and one of the most robust among these is a form of speciation.

(*c*) *Simulation*

The simplest way to illustrate these phenomena is by simulation. Analytical explanations can also be given on the basis of general principles of symmetric bifurcation theory (Cohen & Stewart 2000; Golubitsky & Stewart 2002*a*; Stewart *et al.* 2003). Suppose for simplicity that $k = 1$, and write x_i in place of x^i. The most widely studied family of models is the simplest family that can describe local symmetry-breaking bifurcations, and it is a polynomial equation with terms of degree 3 or less. Technically, it is the general cubic truncation of a centre manifold reduction (Carr 1981) of such a bifurcation. It takes the form

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = \lambda x_i + B(Nx_i^2 - \pi_2) + C(Nx_i^3 - \pi_3) + Dx_i\pi_2 \tag{2.3}$$

for $i = 1, \dots, N$. Here $\lambda, B, C, D \in \mathbb{R}$ are parameters, and

$$\pi_2 = x_1^2 + \cdots + x_N^2,$$

$$\pi_3 = x_1^3 + \cdots + x_N^3.$$

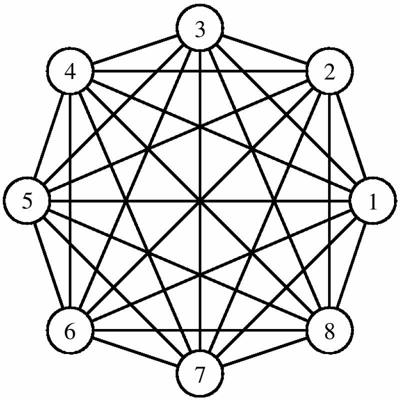

Figure 3. All-to-all coupled-cell network with eight cells (PODs).

Moreover, the variables are *normalized* by translating the origin so that the mean is zero. That is,

$$\pi_1 = x_1 + \cdots + x_N \equiv 0.$$

This mathematically convenient step means that the x^i represent the *deviation* from the mean phenotype, rather than the phenotype itself. Observe that these equations possess the two symmetry features listed above.

Mathematically, such a model constitutes a *coupled-cell system*, a network of interacting ODEs. Each cell (node) of the network has its own *internal dynamic*, and the edges of the network represent *couplings* between nodes. See Golubitsky & Stewart (2002a, b) for the theory of such systems. Here the network is *all-to-all coupled*, meaning that every cell interacts with every other cell. The symmetry constraint implies that all cells have the same internal dynamic (all PODs obey the same rules for 'self') and all couplings are identical (each POD interacts in the same way with all non-self PODs). Figure 3 illustrates the structure of such a network in the case of eight PODs.

Figure 4 shows typical time-series of 25 PODs obeying (2.3), all plotted on the same axis, with λ increasing slowly from negative values to positive ones. Initially, all 25 observations are identical; as λ passes through zero there is a jump bifurcation, in which six PODs take identical negative values of x^j and 19 take identical positive values. Bifurcations from a uniform state (all x^j equal) to a state with several clumps (usually two) are typical of all $\boldsymbol{S}_N$-equivariant dynamical systems, for symmetry reasons (see Golubitsky & Stewart 2002a).

(*d*) *Universal features*

Cohen & Stewart (2000) analyse models like (2.1), with symmetry (2.2), using standard methods of symmetric bifurcation theory. They show that *independently of the form of* f, any symmetry-breaking steady-state bifurcation (if it occurs) will exhibit several universal features.

(i) Generically, the divergence occurs by jump bifurcation: 'punctuated equilibrium' in the terminology of evolutionary biology (Eldredge & Gould 1972).

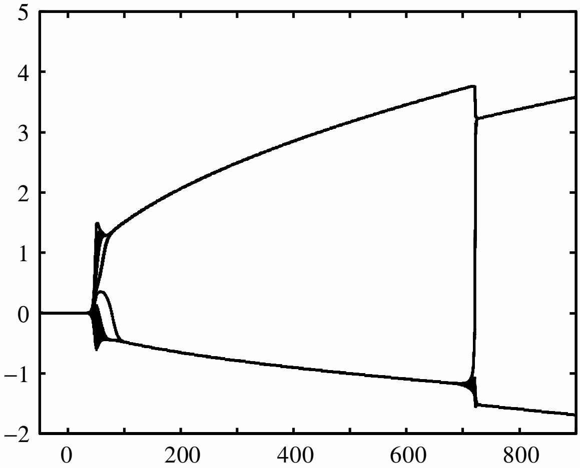

Figure 4. Symmetry-breaking bifurcation in the model (2.3) with $N = 25$ PODs ($C = -1$, $D = -0.2$). Time series of all cells are superposed, with λ horizontal and x_j vertical for each j. Note the 'primary' bifurcation to a state in which 19 PODs take positive and equal values, and the other six take negative and equal values. There is also a 'secondary' bifurcation between primary branches when $\lambda \sim 720$.

(ii) The initial bifurcation is most likely to lead to the formation of *two* distinct morphs or species. Technically, primary bifurcations can occur stably to 2-morph states. Secondary bifurcations to 3-morph states will be far less common, and in the cubic truncation (2.3) all 3-morph states turn out to be unstable (Dias & Stewart 2003).

(iii) Despite the discontinuity involved in a jump bifurcation, mean phenotypes change *smoothly* during such a speciation event; so, arguably, do mean-field genotypes related to continuous characters.

Mathematically, the most counter-intuitive implication of these models is the conjunction of items (i) and (iii): jumps that leave the mean essentially unchanged. However, this conclusion is biologically quite plausible: the bifurcating species adopt phenotypes that differ from the original in *opposite* ways, which is likely to be a sensible way to exploit resources more effectively. For example, suppose that the original species is a finch, and the phenotypic variable is beak length. If some birds develop larger beaks, while others develop smaller ones, then the two species between them can exploit a wider range of seed sizes. It is less clear what advantage can accrue if some birds develop larger beaks than before, while the rest develop beaks that are larger still.

(*e*) *Related observations*

There is some evidence to support the above predictions. Similar phenomena occur in theoretical analyses, based on different modelling hypotheses from ours. Thus, Higashi *et al.* (1999) argue that sympatric speciation can be accomplished through sexual selection without disruptive natural selection. They study how the probability distributions of female preference and male phenotype co-evolve, finding that each splits into two groups, diverging from the original mean in opposite directions. Similarly, Dieckmann & Doebeli (1999) develop a model with multilocus genetics and assortative mating. Their simulations show that new branches diverge from the mean in opposite directions. Another example occurs in the adaptive dynamics models of Kisdi & Geritz (1999).

What about observational evidence? The process of speciation cannot be observed directly on very long time-scales, although some information can be deduced from the fossil record. Direct observations of speciation can sometimes be made on a shorter time-scale: examples include Darwin's finches. As a proxy for evolution, it is sometimes possible to observe two related species or subspecies coexisting in a given environment. Their phenotypes then may be expected to correspond to a 'sympatric' context, before speciation has occurred. (Phenotypic divergences here are normally interpreted as character displacement (Salthe 1972), but they can also be seen as a surrogate for the early stages of small-scale divergences of species.) We can compare these phenotypes with those that occur when only one of the species exists in a given location (ideally with much the same environment as for the sympatric case). This second situation resembles what would happen after speciation has occurred.

With this interpretation, the 'constant mean' prediction is consistent with several field studies, normally interpreted as evidence for character displacement. These include Beauchamp & Ullyett (1932), who study temperature preference in the flatworm *Planaria gonocephala*; Huey *et al.* (1974, 2000) and Huey & Pianka (1974) on the sizes of the skinks *Typhlosaurus lineatus* and *T. gariepensis*; work of Bantock & Bayley (1973) and Bantock *et al.* (1975) on shell sizes in the snails *Cepea nemoralis* and *C. hortensis*; and investigations by Fenchel (1975*a*, *b*) on the mud-snails *Hydrobia ulvae* and *H. ventrosa* (but note that Barnes (1993, 1994, 1996) expresses reservations about those results). See Stewart *et al.* (2003) for further discussion.

By a neat historical irony, there is evidence of a similar effect in Darwin's finches, with the phenotypic variable being beak size. The species *Geospiza fortis* and *G. fuliginosa* occur in both sympatric and allopatric populations. *G. fortis* is allopatric on Daphne, and *G. fuliginosa* is allopatric on Crossman. Moreover, the two species are sympatric on a number of islands, which occur in three groups: Abingdon, Bindloe, James, Jervis; Albemarle, Indefatigable; and Charles, Chatham. Figure 5, adapted from Lack (1968), shows the differences in beak size between these species on these groups of islands. The mean beak sizes of *G. fortis* and *G. fuliginosa* are *ca.* 10 mm in the allopatric case. In all three sympatric populations, the mean for *G. fortis* is *ca.* 12 mm, while that for *G. fuliginosa* is *ca.* 8 mm. Thus they diverge in opposite directions with a constant mean of *ca.* 10 mm.

3. Fitness functions

Our model casts some new light on the classical concept of a 'fitness landscape', a concept that goes back to Wright (1931, 1932), who used the term 'peaks of adaptedness', and did not formulate a mathematical description. In the standard mathematical formulation, the fitness of a genotype (or phenotype) x is measured by some real-valued function $\phi(x)$, and evolution maximizes $\phi(x)$. The usual model is that x 'climbs' the gradient of ϕ, by obeying an ODE

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \nabla\phi,$$

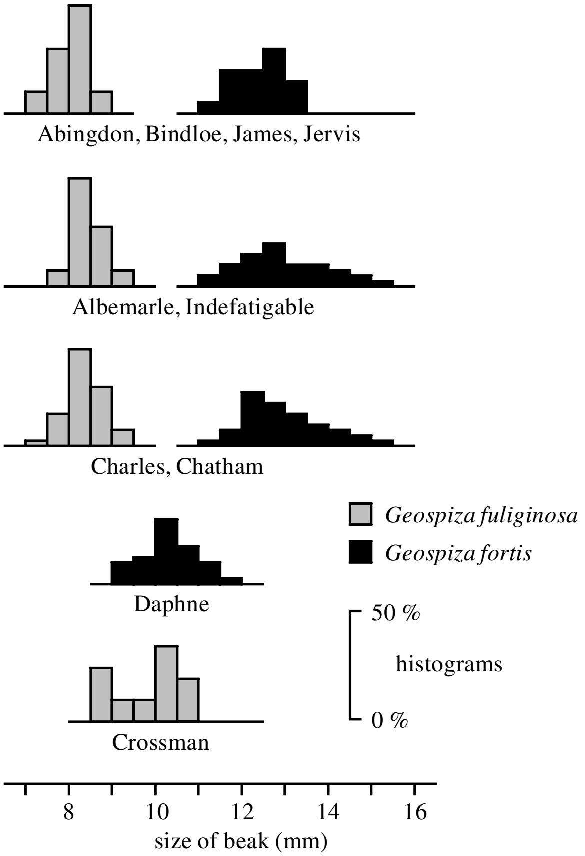

Figure 5. Beak sizes in allopatric and sympatric populations of *Geospiza* in the Galápagos Islands. (After Lack (1968).)

where ∇ indicates the gradient. The graph of ϕ is the metaphorical landscape referred to here. The conventional assumption (see ch. 2 of Kauffman (1993) for extensive references) is that all organisms in the same population climb the same landscape, which effectively assumes that their evolutionary paths are decoupled from each other. This implies that all organisms must adopt the same strategy, give or take some random noise. This assumption effectively makes symmetry breaking impossible, so speciation becomes very puzzling from this point of view. The literature at the time of Wright expresses this puzzlement at length.

In practice, it is unreasonable to assume that all organisms in the same population have the same fitness function. Indeed, a major factor in the fitness of any organism is the behaviour of the other organisms in its species, with whom it is most closely competing. The usual interpretation of Wright's concept effectively treats evolution like a soccer match in which each player has its own ball and completely ignores all the other players.

There is a more reasonable concept of 'fitness landscape', which assigns a numerical value to the collective fitness of the distribution of phenotypes within the population. Such a function arises naturally in the system (2.3). Specifically, it so happens (Dias & Stewart 2003) that (2.3) is a gradient flow. That is, there exists a function

$G : \mathbb{R}^{N-1} \to \mathbb{R}$ for which the ODE takes the form

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = \frac{\partial}{\partial x_i} G(x, \lambda), \tag{3.1}$$

or, more compactly,

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \nabla G(x, \lambda). \tag{3.2}$$

Here

$$G(x) = \tfrac{1}{2}\lambda\pi_2 + B(\tfrac{1}{3}N\pi_3) + C(\tfrac{1}{4}N\pi_4) + \tfrac{1}{4}D\pi_2^2,$$

bearing in mind that $\pi_1 = x_1 + \cdots + x_N = 0$. Here π_2, π_3 are as before, and $\pi_4 = x_1^4 + \cdots + x_N^4$.

Such systems flow 'uphill' to local maxima of G. So we can interpret G as a type of *fitness function*, and the graph of G as a *fitness landscape*. The evolutionary dynamic moves the system in the direction of locally increasing fitness.

Unlike classical fitness functions, however, this G depends not on just one organism, but on them all. The 'fitness' of a phenotype x^i in POD i is not just a function of x^i, but of all the other x^j as well. This statement makes good biological sense, and is crucial to collective effects in populations of organisms, but it is largely ignored by classical genetics.

Even though the fitness function provides a collective measure, the gradient ODE (3.1) states that each POD climbs its own fitness slope, *as if the other phenotypes were held constant* at any given instant. So each POD behaves 'selfishly', to improve its own fitness. However, *all PODs* are doing this *simultaneously*, and, as a result, the entire population evolves towards a local maximum of the collective fitness.

Note that we did not put the above observations into the model as an assumption. They are consequences of the gradient structure, which is a theorem (for the cubic truncation, at any rate), not a modelling hypothesis.

The collective effect of these individual actions leads to a simple form of self-organization, in which a selfish rule for individual improvement leads to a collective evolution towards a state of (locally) maximal fitness for the entire population: a *fitness peak*. Maximization of fitness is a conventional metaphor for natural selection. The question is fitness *of what*? In the current model, it is neither the fitness of a single organism nor that of the entire population. It is the collective fitnesses of the individual organisms of the population, each being considered in the context of what the others are doing.

The distinction is subtle, but it places old debates about 'species selection' in a new light. The debates of the 1950s concluded that natural selection acts on organisms, not on species. (This removed one possible explanation of speciation as the appearance of a new fitness peak in the species landscape.) However, the terms of the debate assumed that the alternatives were selection of an organism in competition with another organism, or selection of a species in competition with another species. Equation (3.2) represents a different kind of 'species selection'. The only direct competition is between organisms. But because the fitness of an organism depends on the phenotypes of the other organisms, the entire species evolves collectively. So the old conclusion that evolution 'really' acts on organisms holds good, but the deduction that species cannot compete with other species, following a collective strategy rather than an individual strategy, does not. The contextual effects of this kind of 'fitness' bind the individuals together strategically. This binding does not depend on conscious awareness of the collective behaviour, or even of the existence of the collective: it stems ultimately from 'selfish' interactions between individuals. Note that this process does not involve 'cooperation'. Organisms do not have to be *aware* of what the others are doing. They just have to interact with them ecologically.

(*a*) *How symmetry breaks*

The geometry of generic bifurcation in an $\boldsymbol{S}_N$-symmetric system differs considerably from what we would expect in a general asymmetric system. There, the generic change would be from a local maximum to a saddle, as an eigenvalue changes sign. But in the $\boldsymbol{S}_N$-symmetric case, symmetry implies that the bifurcation from one species to two occurs when the local fitness peak changes from a local maximum to a local minimum. This change is generic thanks to symmetry, which here forbids the occurrence of saddle points near the origin (see Stewart *et al*. 2003). In the absence of symmetry, such a change would be highly non-generic.

This is what makes it possible for symmetry to break, so that the population can speciate. If each POD climbed the 'same' fitness landscape, independently of the other PODs, and they all started from the same place, then they would all ascend the same fitness peak, and the state would remain monomorphic. However, if each POD climbs 'the same' fitness landscape, in the context of the other PODs, then symmetry can break, and some PODs can behave differently from the others because (thanks to small, 'random' perturbations in initial conditions just after bifurcation) they exist in a different context. The system maximizes its collective fitness, and this may involve some PODs taking up different phenotypes from others.

For example, suppose that the population consists of birds, and x^i is beak size for POD i. (We temporarily abandon the convention that all phenotypes are normalized to have zero mean, for simplicity of description.) Initially, all PODs are at the same value, say $x^i = 10$ mm for all i. Then, as environmental parameters λ pass through some critical value, this state changes from a local maximum to a local minimum of G. Now each POD potentially can improve its fitness by changing its phenotype.

If all PODs change in the same direction, however, fitness is *not* improved. There is now *no* fitness peak, with all x^i equal. However, there are numerous nearby fitness peaks with unequal x^i, including (for group-theoretic reasons) many in which the phenotypes x^i take just two distinct values. As the x^j fluctuate at random, the flow of the ODE takes the state rapidly (Cohen & Stewart 2000) to one in which some PODs take on values larger than 10 mm, and others take on values smaller than 10 mm. (Indeed, since being equal to 10 mm has probability zero, it must be one or the other for each POD.)

This initial random bias improves the collective fitness (because *any* local move away from equality does so), and is therefore amplified by the flow of the ODE. Because of the symmetry of a species, the population collectively does better after *any* random change. As evolution progresses, the population collectively climbs the gradient of G, which reinforces the split, and we soon find that some subset of PODs has settled for a higher value, say 12 mm, while the rest have settled for a lower value, say 8 mm. That there are just two values follows from the analysis of Cohen & Stewart (2000), Stewart *et al*. (2003) and Dias & Stewart (2003). And the same analysis explains why we expect to find one value lower than 10 mm and the other higher.

The biological interpretation is that under appropriate circumstances (namely, instability of the single-species state) random changes in phenotype automatically improve the collective fitness and trigger a bifurcation to a polymorphic state. The population as a whole does better if the individuals adopt phenotypes that allow some birds to specialize on different resources from the others—say some prefer small seeds while others prefer large ones. This process does not require any kind of 'altruism', though collective effects of the kind that occur in this example could provide a new explanation of altruism. It also does not require any conscious awareness of what the other birds are doing; it is a consequence of individuals following 'selfish' rules subject to having to interact with the rest of the population (for example, by competing for the same resources). The birds do not even have to *see* or *encounter* each other: a bird cannot eat seeds that have already been consumed by an earlier bird, and it does not need to have any knowledge of that bird's existence for this to be so.

So, with each POD selfishly evolving towards what suits *it* best, in the context that it happens to be experiencing, the small random fluctuations that trigger the instability lead to a coherent, collective, *self-organized* movement of the population to a polymorphic state. This movement is constrained by the symmetry of the dynamic in phase space.

4. Removing exact symmetry assumptions

We now begin to make the model more 'realistic' by changing its structure. One clear idealization in the analysis so far is the occurrence of exact symmetry. In reality, organisms in a given species are not identical. There are several ways to deal with this objection. The main modelling point is that, mathematically, the best way to model an approximately symmetric system is to view it as a small perturbation of an idealized system with exact symmetry. The main phenomena in the perturbed system are perturbed versions of what we expect in the idealized system, but the idealized system is much easier to analyse and to understand.

To biologists, this (entirely reasonable) argument often sounds like special pleading, so here we take a more direct route. We can add stochastic terms, change the general equations (2.1) or the special case (2.3) to break the symmetry, or both. We now show, by numerical experiment and theoretical discussion, that the main phenomena associated with (2.3) survive such modifications, so the quest for greater biological realism does not alter the main conclusions derived from the less realistic, but far more tractable, equations (2.3).

(*a*) *Stochastic ODEs*

One way to introduce variability within a given species is to convert (2.1) into a stochastic model

$$\mathrm{d}x = F(x,\lambda)\,\mathrm{d}t + \sigma\,\mathrm{d}w,$$

where $\sigma\,\mathrm{d}w$ indicates Gaussian random noise (Brownian motion) with standard deviation σ. This represents the inherent variability of phenotype within the same species and prevents individuals from being exactly the same (see Itô 1961; McKean 1969; Gihman & Skorohod 1970).

In practice we discretize and iterate:

$$x(t+1) = x(t) + \epsilon F(x,\lambda) + \sigma\rho.$$

Figure 6. (*a*) Stochastic model also leads to speciation; (*b*) more noise makes the clumps larger, but they remain clumps.

Here ϵ is small (usually $\epsilon = 0.0001$), σ determines the size of the noise, and ρ is a random variable distributed uniformly between -0.5 and 0.5. (Uniformity is assumed for ease of simulation and is an unimportant side issue here.) Figure 6*a* shows a typical simulation. The system still breaks symmetry, leading to speciation, and the three predictions remain valid. The main difference is that, the larger the noise level, the broader the two clumps in phenotypic space become (see figure 6*b*). (Of course, very large σ would cause the clumps to merge into one.)

In fact there is a general theoretical reason for expecting the effects of these changes to the model to be small. The long-term statistical behaviour of stochastic nonlinear ODEs

$$\mathrm{d}x = F(x, a)\,\mathrm{d}t + \sigma\,\mathrm{d}w \tag{4.1}$$

is governed by the *Fokker–Planck equation*

$$\frac{\partial u}{\partial t} = \sigma\nabla^2 u - \nabla \cdot (uF).$$

Here the function $u = u(t)$ represents a time-varying smooth probability density, the term $\sigma\nabla^2 u$ represents random diffusion, and $\nabla \cdot (uF)$ represents the deterministic flow ($\sigma = 0$). Intuitively, solutions of (4.1) are like solutions of the deterministic equation subjected to random kicks at all instants of time, so for low noise ($\sigma \ll 1$) they look like slightly irregular versions of deterministic solutions. Formalizing these ideas requires substantial technical effort (again see Itô 1961; McKean 1969; Gihman & Skorohod 1970).

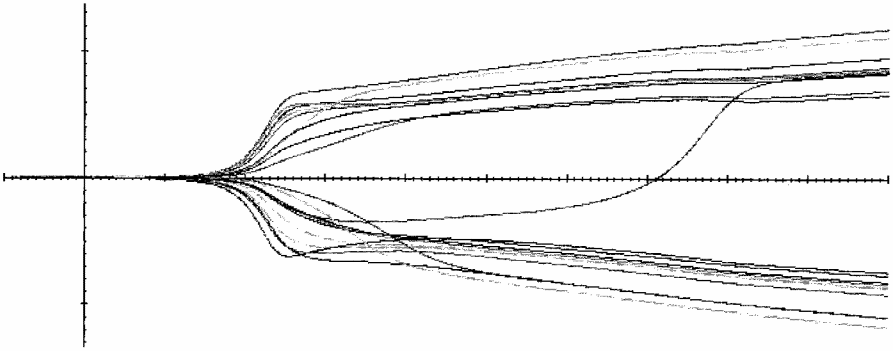

Figure 7. Imposed broken symmetry.

It is known that as t tends to infinity, any solution of the Fokker–Planck equation converges to a unique steady state U, called the *stationary probability density function.* This implies that if F has symmetry group Γ, and the noise is Γ-invariant, then U must be Γ-invariant. In particular if, for specific parameter values a, the attractors of the deterministic system are hyperbolic equilibria and $\sigma \ll 1$, then U has peaks near all stable equilibria and is small everywhere else (see Kifer 1986; Zeeman 1988).

As it happens, the gradient dynamic here allows us to describe explicitly how the system behaves in the stochastic case. We have a multivariate stochastic ODE of the form

$$\mathrm{d}x = \nabla G(x, \lambda)\,\mathrm{d}t + \sigma\,\mathrm{d}w,$$

with w representing 'white noise'. Assuming that w is isotropic—a reasonable modelling hypothesis and one that reflects biologists' standard assumption that phenotypic and genetic variation are 'random'—we can solve the Fokker–Planck equation explicitly to find the stationary PDF, which takes the form

$$P(x) = k\mathrm{e}^{G(x,\lambda)/\sigma^2},$$

where k normalizes the total integral to unity. This distribution has fitness peaks at the local maxima of G, and if those maxima are non-degenerate, then locally the shape is Gaussian. In traditional terminology (Salthe 1972) Gaussian *niches* appear. But these niches are not predetermined by the environment: they arise through the collective interactions of the organisms, in the context of the environment.

(*b*) *Forced symmetry breaking*

Alternatively, we can remain in a deterministic setting, but break the $\boldsymbol{S}_N$ symmetry of (2.3) by making the coefficients A, B, C, D vary slightly with the index i, and by replacing the terms $x_1 + \cdots + x_N$ and $x_1^2 + \cdots + x_N^2$ by $r_1x_1 + \cdots + r_Nx_N$ and $s_1x_1^2 + \cdots + s_Nx_N^2$, where the r_j and s_j differ slightly from unity. Here 'slightly' is governed by a new parameter g, which is typically 0.1 or thereabouts, indicating a 10% variation of the parameter values. In simulations, these variations are defined at the start of each run using a random number generator.

Now we find that the PODs still speciate, much as before, except that their final state is not as tightly clustered (see figure 7). Speciation persists even when the asymmetry is substantial.

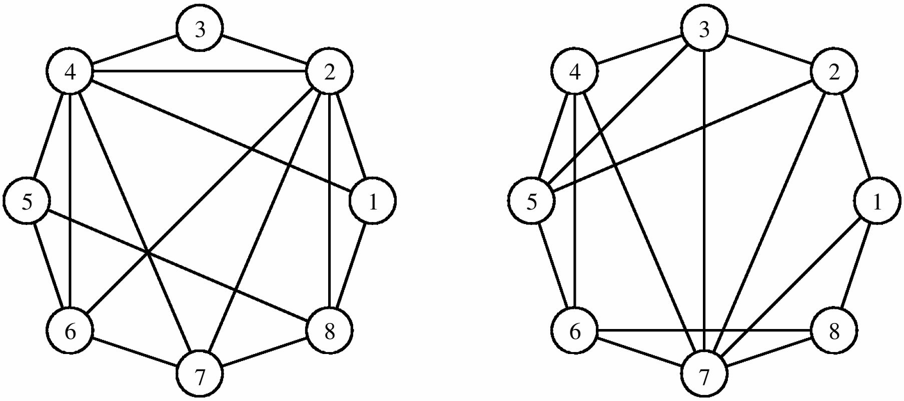

Figure 8. Random networks with eight cells (PODs).

Again, there is a theoretical reason for this resemblance: this time it is *normal hyperbolicity*, which is a generic property in this context (Arrowsmith & Place 1990). Normally hyperbolic equilibria persist under all sufficiently small perturbations, symmetric or not.

5. Random nets and iterated function schemes

We now move on to some new variations on the above theme.

The all-to-all coupled model assumes that all PODs interact with other PODs in exactly the same way. The assumption is not very realistic, and a more reasonable assumption is that at any given time all PODs have the potential to interact with any other POD, but that the interaction may or may not occur. However, if there is an interaction, we continue to assume that it occurs in exactly the same way for all pairs of PODs.

We therefore modify the model, so that at each step in the iteration the network of interactions is chosen by a stochastic procedure. For purposes of illustration we choose a fixed probability p, and retain any given edge of the interaction network with probability p. Figure 8 shows examples of such networks. When iterating the numerical method for POD i we remove from the ODE all terms involving variables x^j that correspond to PODs j that are not currently interacting with POD i. For example, from the term $x_1 + \cdots + x_N$ in the equation for $\dot{x}_i$ we remove all x_j that are not connected to POD i by the current random network.

We can implement the random network approach in two ways. First, we can fix a single random network throughout the simulation. Alternatively, we can choose a new random network at each iteration. The second method is more realistic, but comparison with the first yields some useful insights, so we study both.

Figure 9a shows a typical simulation of this procedure in the case of a fixed network. Again, we see that the PODs speciate into two distinct phenotypes, and aside from minor (though interesting) details, the picture looks very similar to the fully deterministic model. The clusters are not as tight as they are in the symmetric case, but they are still clusters. The changes occurring immediately after bifurcation are complex, because here the system is out of equilibrium, but the system quickly settles down to a more regular state.

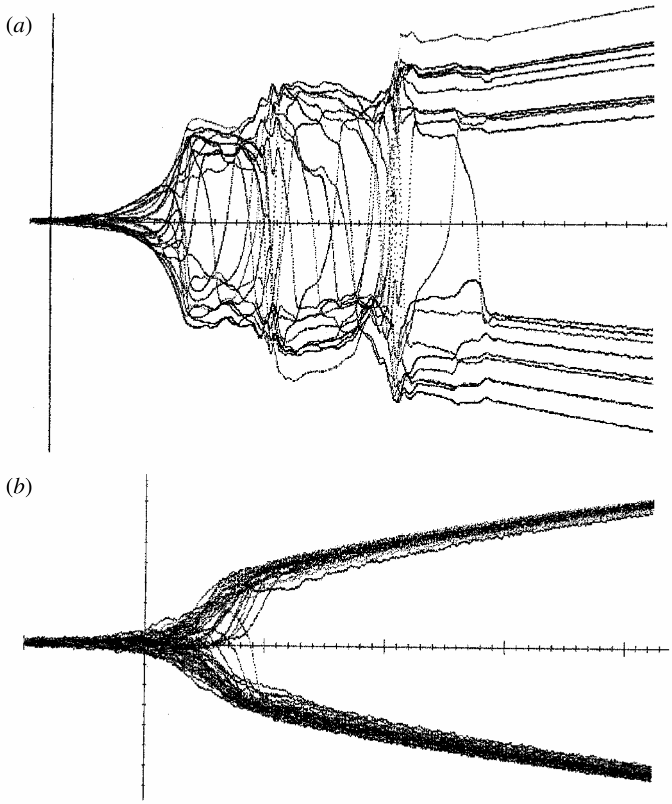

Figure 9. (*a*) Iteration with a fixed random interaction network. (*b*) Iteration with a new random network at each step. Note the tighter clustering in the second case.

If the random network is chosen afresh at each iteration step, as is the case for figure 9*b*, then the clusters become tighter. This model seems especially appropriate to capture the effects of random mating.

The tighter clustering in the second model presumably arises from the *statistical symmetry* of the set of networks. That is, the probability distribution on this set is $\boldsymbol{S}_N$-invariant. The most significant coupling between cells occurs at the linear level, and changing the network randomly at each iteration implies that over a series of iterations the effective coupling is well approximated by the mean coupling over all possible networks. That is, the stochastic model can be approximated by one with mean-field couplings (but retaining nonlinear cell dynamics), and this approximate model has the same form as (2.3) and so behaves in the same manner.

This model has important implications for the role of gene flow, which is usually thought of as a stabilizing influence that prevents speciation (Mayr 1963, 1970). Repeated random selection of an interaction network corresponds to random mating, and acts as a surrogate for gene flow. The effect of gene flow in this sense is indeed stabilizing; however, what it stabilizes is the *two*-species state. In other words, in the presence of selection, gene flow causes individual phenotypic clumps to become tighter, but it does *not* cause them to merge into a single clump.

More formally, we can restate the second model as a nonlinear iterated function scheme (IFS) (Barnsley 1993; Falconer 1990). Each iteration step is a discretization of an ODE

$$\frac{\mathrm{d}x}{\mathrm{d}t} = F^G(x),$$

where G is the current network. Let $\mathcal{G}$ be the set of all possible networks. Then $\{F^G : G \in \mathcal{G}\}$ is a nonlinear IFS. Moreover, the probability distribution induced on $\mathcal{G}$ by the procedure for choosing G is symmetric under $\boldsymbol{S}_N$. Thus, we have a *symmetric iterated function scheme*. In simulations, the dynamics of this IFS seems to break symmetry in the same way that the dynamics of a symmetric ODE would do. There is probably a rigorous explanation of this phenomenon, exploiting the statistical symmetry of the IFS. This looks like an interesting area for new mathematical work. Field & Golubitsky (1992) have developed the beginnings of a theory of symmetric IFSs.

6. Conclusions

We have represented speciation as a process of phenotypic symmetry breaking, in which new species arise when the single-species state becomes unstable. The instability occurs when the organisms concerned can exploit the environment more effectively by adopting a non-uniform strategy; the resulting speciation event is sympatric.

Associated with this type of symmetry breaking are several universal phenomena. The initial bifurcation is a jump bifurcation (punctuated equilibrium), and the typical bifurcation is to a two-species state. Phenotypic means vary smoothly throughout the bifurcation.

In analogues with explicit genetics, allele frequencies should also vary smoothly throughout the bifurcation, for much the same reason. That is, Fisher-style mean-field models are blind to this kind of speciation. Standard deviations, on the other hand, detect such bifurcations reliably: the standard deviation of the entire population is much larger after speciation has occurred. However, the standard deviation for each cluster alone is typically comparable to that for the original single species.

Bifurcations of this kind are highly robust, and most of the structure observed in an idealized, perfectly symmetric model persists in more realistic perturbations of the model. Those perturbations include stochastic noise, forced symmetry breaking in the model equations, and the replacement of a symmetric interaction network by a fixed random network or a sequence of randomly chosen networks.

The models exhibit a simple type of self-organization. However, neither the individual organisms nor the population as a whole self-organizes. Instead, the 'selfish' actions of individuals within the context of the population and the surrounding ecosystem automatically lead to 'organized' collective behaviour, in which speciation can naturally arise in response to changes in environment. One way to formulate this behaviour is through a 'fitness function' that depends on all of the individual strategies, and is thus multidimensional.

The models do not incorporate explicit genetics, but could be modified to do so. The main conclusions would not greatly change, provided the effect of genetics occurs through recombination, rather than being modelled by changes in allele proportions in a mean-field gene pool. Divergence of phenotypes can occur through recombination and selection, without any change in allele proportions, through changes in co-adapted gene complexes (Arthur 1984; Rollo 1995; Salthe 1972; Cohen & Stewart 2000; Stewart *et al*. 2003).

Given that such simple models display such complex and counterintuitive behaviour, it is not surprising that real evolution does too. Verbal models and arguments are too coarse an instrument to resolve these issues. There is room to develop more sophisticated, biologically realistic models. Techniques for doing so already exist.

References

Arrowsmith, D. K. & Place, C. M. 1990 *An introduction to dynamical systems*. Cambridge University Press.

Arthur, W. 1984 *Mechanisms of morphological evolution*. Wiley.

Auyang, S. Y. 1998 *Foundations of complex-system theories*. Cambridge University Press.

Bantock, C. R. & Bayley, J. A. 1973 Visual selection for shell size in *Cepea*. *J. Anim. Ecol.* **42**, 247–261.

Bantock, C. R., Bayley, J. A. & Harvey, P. H. 1975 Simultaneous selective predation on two features of a mixed sibling species population. *Evolution* **29**, 636–649.

Barnes, R. S. K. 1993 Life-history strategies in contrasting populations of the coastal gastropod *Hydrobia*. III. Lagoonal versus intertidal-marine *H. neglecta*. *Vie Milieu* **43**, 73–83.

Barnes, R. S. K. 1994 Investment in eggs in lagoonal *Hydrobia ventrosa* and life-history strategies in north-west European *Hydrobia* species. *J. Mar. Biol. Assoc.* **74**, 637–650.

Barnes, R. S. K. 1996 Breeding, recruitment and survival in a mixed intertidal population of the mudsnails *Hydrobia ulvae* and *H. neglecta*. *J. Mar. Biol. Assoc.* **76**, 1003–1012.

Barnsley, M. F. 1993 *Fractals everywhere*, 2nd edn. Academic.

Beauchamp, R. S. A. & Ullyett, P. 1932 Competitive relationships between certain species of fresh-water triclads. *J. Ecol.* **20**, 200–208.

Carr, J. 1981 *Applications of centre manifold theory*. Springer.

Cohen, J. & Stewart, I. 2000 Polymorphism viewed as phenotypic symmetry-breaking. In *Nonlinear phenomena in physical and biological science* (ed. S. K. Malik), pp. 1–63. New Delhi: Indian National Science Academy.

Darwin, C. 1859 *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*. London: John Murray.

Dias, A. P. S. & Stewart, I. 2003 Secondary bifurcations in systems with all-to-all coupling. *Proc. R. Soc. Lond.* A **459**. (In the press.)

Dieckmann, U. & Doebeli, M. 1999 On the origin of species by sympatric speciation. *Nature* **400**, 354–457.

Eberhard, W. G. 1980 Horned beetles. *Scient. Am.* **242**, 166–182.

Eldredge, N. & Gould, S. J. 1972 Punctuated equilibrium: an alternative to phyletic gradualism. In *Models in palaeobiology* (ed. T. J. M. Schopf). San Francisco, CA: Cooper.

Elmhirst, T. 1998 Symmetry-breaking bifurcations of $\boldsymbol{S}_N$-equivariant vector fields and polymorphism. MSc thesis, Mathematics Institute, University of Warwick, Coventry, UK.

Elmhirst, T. 2001 Symmetry and emergence in polymorphism and sympatric speciation. PhD thesis, Mathematics Institute, University of Warwick, Coventry, UK.

Falconer, K. 1990 *Fractal geometry*. Wiley.

Fenchel, T. 1975*a* Factors determining the distribution patterns of mud snails (*Hydrobiidae*). *Oecologia* **20**, 1–17.

Fenchel, T. 1975*b* Character displacement and coexistence in mud snails. *Oecologia* **20**, 19–32.

Field, M. J. & Golubitsky, M. 1992 *Symmetry in chaos*. Oxford University Press.

Fisher, R. A. 1930 *The genetical theory of natural selection*. Oxford: Clarendon.

Gihman, I. I. & Skorohod, A. V. 1970 *Stochastic differential equations*. Springer.

Golubitsky, M. & Stewart, I. 2002*a* *The symmetry perspective*. Basel: Birkhäuser.

Golubitsky, M. & Stewart, I. 2002*b* Patterns of oscillation in coupled cell systems. In *Geometry, dynamics and mechanics: 60th birthday volume for J. E. Marsden* (ed. P. Holmes, P. Newton, & A. Weinstein), pp. 243–286. Springer.

Golubitsky, M., Stewart, I. & Schaeffer, D. G. 1988 *Singularities and groups in bifurcation theory 2*. Applied Mathematical Science, vol. 69. Springer.

Grant, P. R. 1991 Natural selection and Darwin's finches. *Scient. Am.* **265**, 60–65.

Grant, P. R. & Grant, B. R. 2002 Unpredictable evolution in a 30-year study of Darwin's finches. *Science* **296**, 707–711.

Grant, P. R., Grant, B. R., Smith, J. N. M., Abbott, I. J. & Abbott, L. K. 1976 Darwin's finches: population variation and natural selection. *Proc. Natl Acad. Sci. USA* **73**, 257–261.

Hale, J. & Koçak, H. 1991 *Dynamics and bifurcations*. Springer.

Higashi, M., Takimoto, G. & Yamamura, N. 1999 Sympatric speciation by sexual selection. *Nature* **402**, 523–526.

Huey, R. B. & Pianka, E. R. 1974 Ecological character displacement in a lizard. *Am. Zool.* **14**, 1127–1136.

Huey, R. B., Pianka, E. R., Egan, M. E. & Coons, L. W. 1974 Ecological shifts in sympatry: Kalahari fossorial lizards (*Typhlosaurus*). *Ecology* **55**, 304–316.

Huey, R. B., Gilchrist, G. W., Carlson, M. L., Berrigan, D. & Serra, L. 2000 Rapid evolution of a geographic cline in size in an introduced fly. *Science* **287**, 308–310.

Itô, K. 1961 On stochastic differential equations. Memoirs of the American Mathematical Society, vol. 4. Providence, RI: American Mathematical Society.

Kauffman, S. A. 1993 *The origins of order*. Oxford University Press.

Kauffman, S. A. 1995 *At home in the Universe*. New York: Viking.

Kawecki, T. J. 1997 Sympatric speciation via habitat specialization driven by deleterious mutations. *Evolution* **51**, 1751–1763.

Kifer, Y. 1986 General random perturbations of hyperbolic and expanding transformations. *J. Analyse Math.* **47**, 111–150.

Kisdi, E. & Geritz, S. A. H. 1999 Adaptive dynamics in allele space: evolution of genetic polymorphism by small mutations in a heterogeneous environment. *Evolution* **53**, 993–1008.

Kondrashov, A. S. & Kondrashov, F. A. 1999 Interactions among quantitative traits in the course of sympatric speciation. *Nature* **400**, 351–354.

Lack, D. 1968 *Darwin's finches: an essay on the general biological theory of evolution*. Gloucester, MA: Peter Smith.

McKean, H. P. 1969 *Stochastic integrals*. Academic.

Mainzer, K. 1994 *Thinking in complexity*. Springer.

Maynard Smith, J. 1982 *Evolution and the theory of games*. Cambridge University Press.

Mayr, E. 1963 *Animal species and evolution*. Cambridge, MA: Belknap Press.

Mayr, E. 1970 *Populations, species, and evolution*. Cambridge, MA: Harvard University Press.

Rice, R. & Hostert, E. E. 1993 Laboratory experiments on speciation: what have we learned in 40 years? *Evolution* **47**, 1637–1653.

Rollo, C. D. 1995 *Phenotypes*. London: Chapman and Hall.

Rundle, H. D., Nagel, L., Boughman, J. W. & Schluter, D. 2000 Natural selection and parallel speciation in sympatric sticklebacks. *Science* **287**, 306–308.

Salthe, S. N. 1972 *Evolutionary biology*. New York: Holt, Rinehart and Winston.

Seydel, R. 1994 *Practical bifurcation and stability analysis*. Springer.

Stewart, I., Elmhirst, T. & Cohen, J. 2003 Symmetry-breaking as an origin of species. In *Proc. Conf. Bifurcations, Symmetry, Patterns, Porto, 2000*. Basel: Birkhäuser.

Tregenza, T. & Butlin, R. K. 1999 Speciation without isolation. *Nature* **400**, 311–312.

Vincent, T. L. & Vincent, T. L. S. 2000 Evolution and control system design. *IEEE Control Syst. Mag.* **October**, 20–35.

Wolffe, A. P. & Matzke, M. A. 1999 Epigenetics: regulation through repression. *Science* **286**, 481–486.

Wright, S. 1931 Evolution in Mendelian populations. *Genetics* **16**, 97–159.

Wright, S. 1932 The roles of mutation, inbreeding, crossbreeding and selection in evolution. *Proc. 6th Int. Congr. Genetics* **1**, 356–366.

Zeeman, E. C. 1988 Stability of dynamical systems. *Nonlinearity* **1**, 115–155.